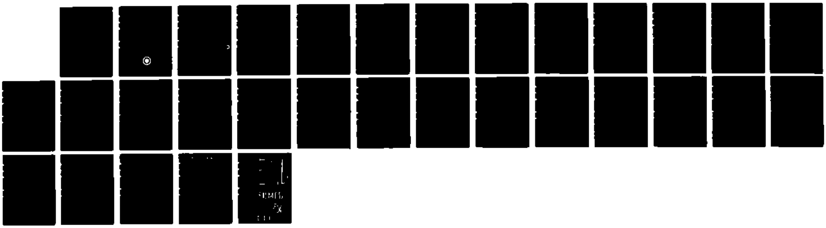
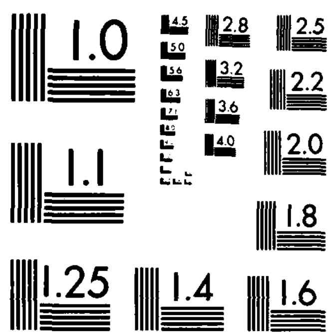
AD-A164 211 ON DETECTION OF NUMBER OF SIGNALS IN PRESENCE OF WHITE NOISE(U) PITTSBURGH UNIV PA CENTER FOR MULTIVARIATE ANALYSIS L C ZHAO ET AL. OCT 85 TR-85-37 1/1

UNCLASSIFIED AFOSR-TR-86-0039 N00014-85-K-0292 F/G 9/4 NL





MICROCOPY RESOLUTION TEST CHART

DARDS-1963-A

AFOSR-TR- 86-0039

(2)

ON DETECTION OF NUMBER OF SIGNALS IN PRESENCE OF WHITE NOISE*

L. C. Zhao

P. R. Krishnaiah

Z. D. Bai

Center for Multivariate Analysis
University of Pittsburgh

Center for Multivariate Analysis

University of Pittsburgh

ON DETECTION OF NUMBER OF SIGNALS IN PRESENCE OF WHITE NOISE*

L. C. Zhao

P. R. Krishnaiah

Z. D. Bai

Center for Multivariate Analysis
University of Pittsburgh

October 1985

Technical Report No. 85-37

Center for Multivariate Analysis
Fifth Floor, Thackeray Hall
Univeristy of Pittsburgh
Pittsburgh, PA 15260

*This work is supported by Contract N0001485K0292 of the Office of Naval Research and Contract F49620-85-C-0008 of the Air Force Office of Scientific Research. The United States Government is authorized to reproduce and distribute reprints for governmental purposes notwithstanding any copyright notation hereon.

1. INTRODUCTION

In the area of signal processing, it is of interest to detect the number of signals in presence of noise and estimate the parameters of the signals. The problem of estimation of the number of signals was discussed by Liggett (1973), Wax, Shan and Kailath (1984) and others in the literature within the framework of testing for the equality of the last few eigenvalues of the covariance matrix. The model considered by them involve expressing the observation vector as the sum of Gaussian white noise vector and a vector of certain linear combinations of (random) signals radiated by sources. In this case, the number of signals is related to the multiplicity of the smallest eigenvalues of the covariance matrix of the observation vector. The problem of testing the hypothesis of the multiplicity of the smallest eigenvalues of the covariance matrix was dealt extensively in multivariate statistical literature (e.g., see Anderson (1963), Krishnaiah (1976), and Rao (1983). Wax and Kailath (1984) considered the problem of determination of the number of signals using information theoretic criteria proposed by Akaike (1972), Rissanen (1978) and Schwartz (1978).

In the present paper, we use an alternative information theoretic criterion for detection of the number of signals and establish its consistency. In Section 2 of the paper, we state briefly the problems considered in this paper. In Sections 3 and 4, we establish the consistency of our procedures when the variance of the white noise is unknown and known respectively. In the above sections, we assumed that the distribution underlying the observations is complex multivariate normal. In Section 5, we establish the consistency of our procedure when the variance of white noise is unknown and satisfies certain condition and the underlying distribution is complex elliptically symmetric. The problem of detection of the number of signals when the noise covariance matrix is arbi-

Errata Sheet

Correction to "On Detection of Number of Signals in Presence of Noise" by L. C. Zhao, P. R. Krishnaiah and Z. D. Bai, CMA Technical Report 85-37

Page 11, lines 11 - 12: Replace

"But, for $k > q$, the above difference is not asymptotically positive with probability one"

With:

"But, for $k > q$, $\log L_k - \log L_q$ is <u>not</u> distributed asymptotically as chi-square".

trary is discussed by the authors in a forthcoming paper when an independent estimate of the above covariance matrix is available. In this case, the problem is equivalent to finding the multiplicity of the smallest eigenvalues of $\Sigma_1\Sigma_2^{-1}$ where Σ_1 and Σ_2 respectively denote the covariance matrices of the observation vector and noise vector. This problem is related to the problem of relationship between two covariance matrices studied by Rao (1983). In another forthcoming paper, we considered the problem of drawing inference on the rank of the regression matrix using information theoretic criteria similar to those used in the present paper.

2. PRELIMINARIES AND STATEMENT OF PROBLEMS

Consider the model

$$\underset{\sim}{x}(t) = A\underset{\sim}{s}(t) + \underset{\sim}{n}(t) \tag{2.1}$$

where $A = [A(\Phi_1),\ldots,A(\Phi_q)]$, $\underset{\sim}{s}(t) = (s_1(t),\ldots,s_q(t))'$, $\underset{\sim}{n}(t) = (n_1(t),\ldots,n_p(t))'$ and $q < p$. In the above model, $\underset{\sim}{n}(t)$ is the noise vector distributed independent of $\underset{\sim}{s}(t)$ as complex multivariate normal with mean vector $\underset{\sim}{0}$ and covariance matrix $\sigma^2 I_p$. Also, $\underset{\sim}{s}(t)$ is distributed as complex multivariate normal with mean vector $\underset{\sim}{0}$ and nonsingular covariance matrix Ψ and $A(\Phi_i)$: $p \times 1$ is a complex vector of functions of the elements of unknown vector $\underset{\sim}{\Phi}_i$ associated with i-th signal. Also, $s_i(t)$ is the waveform associated with i-th signal. Then, the covariance matrix Σ of $\underset{\sim}{x}(t)$ is given by

$$\Sigma = A\Psi\bar{A}' + \sigma^2 I \tag{2.2}$$

where $\bar{A}'$ denotes the transpose of the complex conjugate of A. We assume that $\underset{\sim}{x}(t_1),\ldots,\underset{\sim}{x}(t_n)$ are independent observations on $\underset{\sim}{x}(t)$. Now, let $\lambda_1 \geq \ldots \geq \lambda_p$ denote the eigenvalues of Σ, and $\theta_1 \geq \ldots \geq \theta_q$ denote the nonzero eigenvalues of $A\Psi\bar{A}'$. Also, let H_q denote the hypothesis $\lambda_q > \lambda_{q+1} = \ldots = \lambda_p = \sigma^2$. Under H_q, $\lambda_i = \sigma^2 + \theta_i$ $(i = 1, 2,\ldots,q)$ and $\lambda_{q+j} = \sigma^2 (j=1,2,\ldots,p-q)$. So, H_q is equivalent to the hypothesis that q signals are transmitted. Various procedures (e.g., see Anderson (1963), and Krishnaiah and Waikar (1971, 1972)) are available in the literature for testing the hypothesis H_q for given value of q. Wax and Kailath (1984) used Akaike's AIC criterion and Schwartz-Rissanen minimum distance length (MDL) criterion for model selection for determination of the value of q. According to the AIC criterion, the value of q is estimated to be $\hat{q}$ where $\hat{q}$ is

chosen such that

$$AIC(\hat{q}) = \min\{AIC(0),\ldots,AIC(p-1)\} \tag{2.3}$$

and

$$AIC(k) = -2\log L_k + 2\nu(k,p). \tag{2.4}$$

L_k is the likelihood ratio test statistic for testing H_k against the alternative that Σ is arbitrary, and $\nu(k,p)$ denotes the number of free parameters that have to be estimated under H_k. According to the MDL criterion, the value of q is estimated as $\hat{q}$ where $\hat{q}$ is chosen such that

$$MDL(\hat{q}) = \min\{MDL(0),\ldots,MDL(p-1)\} \tag{2.5}$$

$$MDL(k) = -\log L_k + \frac{\log N}{2}\nu(k,p). \tag{2.6}$$

In the present paper, we consider the following alternative information theoretic criterion for model selection for estimation of the value of q. According to this new information theoretic criterion for model selection, we estimate q with $\hat{q}$ where $\hat{q}$ is chosen such that

$$I(\hat{q},C_N) = \min\{I(0,C_N),\ldots,I(p-1,C_N)\} \tag{2.7}$$

$$I(k,C_N) = -\log L_k + C_N\,\nu(k,p) \tag{2.8}$$

and C_N is chosen such that

$$\lim_{N\to\infty}\{C_N/N\} = 0 \tag{2.9}$$

$$\lim_{N\to\infty}\{C_N/\log\log N\} = \infty. \tag{2.10}$$

We are interested in establishing the strong consistency of the above procedure for the cases when σ^2 is unknown and known under the assumption that the distribution underlying the data is complex multivariate normal. We are also interested in extending the above results to the situation when the underlying distribution is complex elliptically symmetric. The probability of correct detection of the procedure proposed by us is given by

$$P(CD) = P[I(q,C_N) - I(k,C_N) < 0; \quad k = 0,1,\ldots,(p-1);\ k \neq q \mid H_q].$$

Investigation has to be made on the evaluation of P(CD).

3. CONSISTENCY OF $I(\hat{q},C_N)$ CRITERION WHEN σ^2 IS KNOWN AND THE UNDERLYING DISTRIBUTION IS COMPLEX MULTIVARIATE NORMAL

In this section, we establish the consistency of the estimate $\hat{q}$ of q when the criterion $I(\hat{q},C_N)$ is used and σ^2 is unknown. The main result of this section is stated in the following theorem:

<u>THEOREM 3.1.</u> Suppose $\underset{\sim}{x}(t)$ is a complex, stationary process with $E(\underset{\sim}{x}(t)) = \underset{\sim}{0}$ and $E(\underset{\sim}{\bar{x}}'(t)\underset{\sim}{x}(t))^{2+\varepsilon} < \infty$ for some $\varepsilon > 0$. Also, we assume that $\{\underset{\sim}{x}(t_i), i=1,2,\dots\}$ is a stationary and ϕ - mixing sample sequence with ϕ being decreasing and $\sum_{n=1}^{\infty} \phi^{\frac{1}{2}}(n) < \infty$. Also, $\delta^2(u,v) > 0$ for $u,v=1,2,\dots,p$, and $Y_i = (y_{iuv}) = \underset{\sim}{x}(t_i)\underset{\sim}{\bar{x}}'(t_i) - \Sigma$, where

$$\delta^2(u,v) = E\, y_{1uv}^2 + 2 \sum_{i=1}^{\infty} E\{y_{1uv}y_{1+i,uv}\}. \tag{3.1}$$

Let $\hat{q}$ be chosen such that

$$I(\hat{q},C_N) = \min\{I(0,C_N),\dots,I(p-1,C_N)\} \tag{3.2}$$

where $I(k,C_N)$ was defined by (2.8) and C_N is chosen satisfying (2.9) and (2. 10). Then $\hat{q}$ is a strongly consistent estimate of q.

We need the following results to prove the above theorem.

<u>LEMMA 3.1</u>. Suppose $\{x_i, i \geq 1\}$ is a stationary ϕ-mixing sequence with $E(x_1) = 0$ and $E(|x_1|^{2+\varepsilon}) < \infty$ for some $\varepsilon > 0$. Also, ϕ is decreasing with $\sum_{n=1}^{\infty} \phi^{\frac{1}{2}}(n) < \infty$. Then

$$\lim_{N\to\infty} \sup\{\sum_{i=1}^{n} x_i/(2n\delta^2 \log\log n\, \delta^2)^{\frac{1}{2}}\} = 1 \quad \text{a.s.} \tag{3.2}$$

where $\delta^2 = EX_1^2 + 2\sum_{i=1}^{\infty} EX_1X_{1+i} \neq 0$ is assumed. Here, we note that $\sum_{n=1}^{\infty} \phi^{\frac{1}{2}}(n) < \infty$ implies $\delta^2 < \infty$.

For a proof of the above lemma, the reader is referred to Reznik (1968) or Stout (1974).

<u>LEMMA 3.2.</u> Suppose that A, A_n, $n=1,2,\ldots$ are all $p\times p$ symmetric matrices such that $A_n - A = O(\alpha_n)$ and $\alpha_n \to 0$ as $n \to \infty$. Denote by $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_p$ and $\lambda_1^{(n)} \geq \ldots \geq \lambda_p^{(n)}$ the eigenvalues of A and A_n respectively. Then we have

$$\lambda_i^{(n)} - \lambda_i = O(\alpha_n) \quad \text{as} \quad n\to\infty,\ i = 1,\ldots,p.$$

<u>PROOF</u>. Without loss of generality, we can assume $A = \mathrm{diag}[\tilde{\lambda}_1 I_{\mu_1},\ldots,\tilde{\lambda}_r I_{\mu_r}]$, where $\tilde{\lambda}_1 > \tilde{\lambda}_2 > \ldots > \tilde{\lambda}_r$. According to Bai (1984), we know $\lambda_i^{(n)} - \lambda_i \to 0$. At first we consider the special case where $r = 1$. For any i,

$$\begin{aligned} 0 &= |\lambda_i^{(n)} I_p - A_n| = |(\lambda_i^{(n)}-\tilde{\lambda}_1) I_p - (A_n-A)| \\ &= (\lambda_i^{(n)}-\tilde{\lambda}_1)^p + \sum_{\ell=1}^{p} (-1)^{\ell} (\lambda_i^{(n)}-\tilde{\lambda}_1)^{p-\ell} D_{\ell}, \end{aligned} \tag{3.3}$$

where D_ℓ *is the sum of all* ℓ*-ordered* principal minors of $A_n - A$. Since $A_n - A = O(\alpha_n)$, we have $D_\ell = O(\alpha_n^{\ell})$. By (3.3) we know

$$|\lambda_i^{(n)} - \tilde{\lambda}_1| \leq \sum_{\ell=1}^{p} |\lambda_i^{(n)} - \tilde{\lambda}_1|^{-\ell+1} O(\alpha_n^{\ell})$$

which implies $|\lambda_i^{(n)} - \tilde{\lambda}_1| = O(\alpha_n)$ as $n\to\infty$ for $i = 1,\ldots,p$.

Now we consider the general case. Suppose $i \leq \mu_1$. We have

$$\begin{aligned} 0 &= \left|\lambda_i^{(n)} I_p - A_n\right| \\ &= \left| \begin{pmatrix} (\lambda_i^{(n)}-\tilde{\lambda}_1) I_{\mu_1} & & & \\ & (\lambda_i^{(n)}-\tilde{\lambda}_2) I_{\mu_2} & & \\ & & \ddots & \\ & & & (\lambda_1^{(n)}-\tilde{\lambda}_r) I_{\mu_r} \end{pmatrix} - (A_n-A) \right| \end{aligned}$$

$$\triangleq \begin{vmatrix} (\lambda_i^{(n)}-\tilde{\lambda}_1)I_{\mu_1}-B_{11}^{(n)} & -B_{12}^{(n)} \\ -B_{21}^{(n)} & B_{22}^{(n)} \end{vmatrix}$$

$$= |B_{22}^{(n)}|\cdot|(\lambda_i^{(n)}-\tilde{\lambda}_1)I_{\mu_1} - B_{11}^{(n)} - B_{12}^{(n)}B_{22}^{(n)-1}B_{21}^{(n)}|.$$

Since $B_{22}^{(n)} \to \text{diag}[(\tilde{\lambda}_1-\tilde{\lambda}_2)I_{\mu_2},\ldots,(\tilde{\lambda}_1-\tilde{\lambda}_r)I_{\mu_r}]$, $B_{22}^{(n)}$ is nonsingular for all large n. Thus

$$|(\lambda_i^{(n)}-\tilde{\lambda}_1)I_{\mu_1} - B_{11}^{(n)} - B_{12}^{(n)}B_{22}^{(n)-1}B_{22}^{(n)}| = 0. \tag{3.4}$$

From $A_n - A = O(\alpha_n)$, it follows that $B_{11}^{(n)} = O(\alpha_n)$ and $B_{12}^{(n)}B_{22}^{(n)-1}B_{21}^{(n)} = O(\alpha_n^2)$. Using the result proved just before, we get

$$\lambda_i^{(n)} - \tilde{\lambda}_1 = O(\alpha_n)$$

for $i = 1, \ldots, \mu_1$. By the same approach, we can prove

$$\lambda_i^{(n)} - \tilde{\lambda}_h = O(\alpha_n), \quad i = \mu_1+\ldots+\mu_{h-1}+1,\ldots,\mu_1+\ldots+\mu_h, \quad h = 1,\ldots,r, \tag{3.5}$$

which complete the proof of the lemma.

Let $\ell_1 \geq \ldots \geq \ell_p$ denote the eigenvalues of $\hat{\Sigma}$, where $N\hat{\Sigma} = \sum_{i=1}^{N} \underset{\sim}{X}(t_i)\bar{\underset{\sim}{X}}'(t_i)$. Using Lemma 3.1 and the conditions imposed on $\{\underset{\sim}{X}(t_i),\ i = 1,2,\ldots\}$, we have

$$\hat{\Sigma} - \Sigma = O\left(\sqrt{\frac{\log\log N}{N}}\right) \quad \text{a.s.} \tag{3.6}$$

Now, applying Lemma 3.2, we obtain

$$\ell_j - \lambda_j = O\left(\sqrt{\frac{\log\log N}{N}}\right) \quad \text{a.s.} \tag{3.7}$$

for $j = 1, 2, \ldots, p$.

When $\underset{\sim}{x}(t)$ is distributed as complex multivariate normal, the likelihood function for testing the hypothesis H_k against the alternative that Σ has general structure is known to be

$$L_k = \left\{ \prod_{i=k+1}^{p} \ell_i^N \Big/ \left(\frac{1}{p-k} \sum_{i=k+1}^{p} \ell_i\right)^{N(p-k)} \right\}. \qquad (3.8)$$

We will first prove the consistency of the method based upon the criterion $I(\hat{q},C_N)$ when $k < q$. Let $G_1(k) = \log L_k$ and

$$G_k = G_1(k) - C_N[k(2p-k)+1] \qquad (3.9)$$

where $k(2p-k)+1$ is the number of free parameters that have to be estimated under the hypothesis H_k and L_k is given by (3.8). Using (3.7), we get

$$\lim_{N\to\infty} \frac{1}{N}(G_1(q)-G_1(k)) = W(q,k) \quad \text{a.s.} \qquad (3.10)$$

where

$$\begin{aligned} W(q,k) &= \log\left(\prod_{i=q+1}^{p} \lambda_i\right) - (p-q)\log\left(\frac{1}{(p-q)} \sum_{i=q+1}^{p} \lambda_i\right) \\ &\quad - \log\left(\prod_{i=k+1}^{p} \lambda_i\right) + (p-k)\log\left(\frac{1}{(p-k)} \sum_{i=k+1}^{p} \lambda_i\right). \qquad (3.11) \\ &= (q-k)\log\left[\frac{1}{(q-k)} \sum_{i=k+1}^{q} \lambda_i \Big/ \left(\prod_{i=k+1}^{q} \lambda_i\right)^{1/(q-k)}\right] \\ &\quad + (p-k)[\log(\alpha_1 A_1+\alpha_2 A_2) - (\alpha_1 \log A_1+\alpha_2 \log A_2)] \end{aligned}$$

where

$$\alpha_1 = (q-k)/(p-k), \qquad \alpha_2 = (p-q)/(p-k)$$

$$A_1 = \frac{1}{(q-k)} \sum_{i=k+1}^{p} \lambda_i, \quad A_2 = \frac{1}{(p-q)} \sum_{i=q+1}^{p} \lambda_i .$$

By the well known arithmetic mean geometric mean inequatity, we have

$$W(q,k) \geq (p-k)[\log(\alpha_1 A_1+\alpha_2 A_2) - (\alpha_1 \log A_1+\alpha_2 \log A_2)]. \tag{3.12}$$

Also, $A_1 > A_2$. By Jensen's inequality, we have

$$W(q,k) > 0. \tag{3.13}$$

Using (3.9), (3.10),(3.13) and $\lim_{N\to\infty}(C_N/N) = 0$, we obtain

$$G(q) - G(k) = NW(q,k)(1+o(1)) \quad \text{a.s.}$$

So, with probability one for large N, we have

$$G(q) > G(k). \tag{3.14}$$

Now we assume $k > q$ and $k \leq p-1$. Without loss of generality we can assume $\sigma^2 = 1$. By (3.7) we have $\lim_{N\to\infty}(\ell_j-1) = 0$ a.s. for $j = q+1,\dots,p$. Using Taylor's expansion, we get for $k > q$

$$G_1(k) = N\{\sum_{i=k+1}^{p}\log(1+\ell_i-1) - (p-k)\log(1+\frac{1}{p-k}\sum_{i=k+1}^{p}(\ell_i-1))\}$$

$$\doteq -\frac{N}{2}\sum_{i=k+1}^{p}(\ell_i-1)^2(1+o(1)) + \frac{N}{2(p-k)}(\sum_{i=k+1}^{p}(\ell_i-1))^2(1+o(1)) \text{ a.s.}$$

By(3.7) we see that

$$\begin{aligned} G_1(k) &= O(\log\log N) \quad \text{a.s., } p-1 \geq k > q \\ G_1(q) &= O(\log\log N) \quad \text{a.s.} \end{aligned} \tag{3.15}$$

From (3.9), (3.15) and $C_N/\log\log N) \to \infty$, we get

$$G(q) - G(k) = C_N(k-q)(2p-k-q) + O(\log\log N)$$
$$= C_N(k-q)(2p-k-q)(1+o(1)) \quad \text{a.s.} \tag{3.16}$$

Thus with probability one for large N we have

$$G(q) > G(k). \tag{3.17}$$

From (3.14) and (3.17), it follows that with probability one for large N

$$\hat{q} = q.$$

Thus the proof of Theorem 1 is completed.

When $\underset{\sim}{x}(t)$ is distributed as real multivariate normal, the proof goes along the same lines as in the complex case.

Wax and Kailath (1985) showed that (MDL(q)-MDL(k)) is asymptotically negative with probability one for $k < q$. But, for $k > q$, *the above difference* is not asymptotically positive with probability one. So, the strong consistency of the MDL criterion does not follow from the arguments of Wax and Kailath (1985). But, it follows from our results by taking $C_N = \frac{1}{2}\log N$. Wax and Kailath (1985) pointed out that the AIC criterion is not consistent. Hannan and Quinn (1979) considered an information theoretic criterion to determine the order of an autoregressive process; this criterion will be discussed in a subsequent communication.

4. DETECTION OF THE NUMBER OF SIGNALS WHEN VARIANCE OF WHITE NOISE IS KNOWN

In Section 3, we discussed a model selection criterion for detection of the number of signals when the distribution underlying the observations is complex multivariate normal and the variance of white noise is unknown. In this section, we derive analogous criterion when the underlying distribution is (real) multivariate normal and the variance of the white noise is known. The strong consistency of the above criterion is also established.

In the model (2.1), we assume that the noise vector $\underset{\sim}{n}(t)$ is distributed as the multivariate normal with mean vector $\underset{\sim}{0}$ and covariance matrix $\sigma^2 I_p$, A is a real matrix of rank $q < N$, and the signal vector $\underset{\sim}{s}(t)$ is distributed independent of $\underset{\sim}{n}(t)$ as a multivariate normal with mean vector $\underset{\sim}{0}$ and nonsingular covariance matrix Ψ. Then, the covariance matrix of $\underset{\sim}{x}(t)$ is $\Sigma = A\Psi\bar{A}' + \sigma^2 I$. We assume that σ^2 is known. Without loss of generality, we assume that $\sigma^2 = 1$. Let $\lambda_1 \geq \ldots \geq \lambda_p$ denote the eigenvalues of Σ. Now, let

$$\Theta_k: \lambda_1 \geq \ldots \geq \lambda_k > \lambda_{k+1} = \ldots = \lambda_p = 1. \tag{4.1}$$

The k-th model M_k is the one for which Θ_k is true. We are interested in selecting one of the p models $M_0, M_1, \ldots, M_{p-1}$.

The likelihood function is given by

$$L(\theta) = -\frac{N}{2}\log|\Sigma| - \frac{N}{2}\operatorname{tr}\Sigma^{-1}\hat{\Sigma} \tag{4.2}$$

where

$$\hat{\Sigma} = \sum_{j=1}^{N} \underset{\sim}{x}_j \underset{\sim}{x}_j'/N. \tag{4.3}$$

Also, let $\delta_1 \geq \ldots \geq \delta_p$ be the eigenvalues of $\hat{\Sigma}$. In addition, let τ denote

the number of δ_i's which are greater than one. Also, let $d \leq \tau$. We will first calculate $L^*(\lambda_{d+1},\ldots,\lambda_p) = \operatorname*{Sup}_{\Phi_d} L(\theta)$ where $\operatorname*{Sup}_{\Phi_d} L(\theta)$ indicates that $L(\theta)$ is maximized subject to the condition that $\lambda_1 \geq \ldots \geq \lambda_d > 1$.

Write $\Lambda = \operatorname{diag}(\lambda_1,\ldots,\lambda_p)$, $\Delta = \operatorname{diag}(\delta_1,\ldots,\delta_p)$. There exist two real orthogonal matrices $\mathcal{O}_1$ and $\mathcal{O}_2$ such that

$$\Sigma = \mathcal{O}_1'\Lambda\mathcal{O}_1, \quad \hat{\Sigma} = \mathcal{O}_2'\Delta\mathcal{O}_2.$$

Put $Q = \mathcal{O}_2\mathcal{O}_1'$. Then we have

$$L(\theta) = -\frac{N}{2}\sum_{j=1}^{p}\log\lambda_j - \frac{N}{2}\operatorname{tr}\Lambda^{-1}Q'\Delta Q.$$

Since Q is orthogonal, we have

$$\operatorname{tr}\Lambda^{-1}Q'\Delta Q \geq \sum_{j=1}^{p}\delta_j/\lambda_j,$$

and the equality holds for $Q = I_p$ (see Von Neumann (1937)). So,

$$\operatorname*{Sup}_{\Phi_d} L(\theta) = \operatorname*{Sup}_{\Phi_d}\{-\frac{N}{2}\sum_{j=1}^{p}\log\lambda_j - \frac{N}{2}\sum_{j=1}^{p}(\delta_j/\lambda_j)\} \tag{4.4}$$

i.e.,

$$L^*(\lambda_{d+1},\ldots,\lambda_p) = -\frac{N}{2}\log(\lambda_{d+1}\cdots\lambda_p) - \frac{N}{2}\sum_{j=d+1}^{p}\delta_j/\lambda_j$$

$$+ \operatorname*{sup}_{\Phi_d}\{-\frac{N}{2}\log(\lambda_1\cdots\lambda_d) - \frac{N}{2}\sum_{j=1}^{d}\delta_j/\lambda_j\} \tag{4.5}$$

$$= -\frac{N}{2}\log(\lambda_{d+1}\cdots\lambda_p) - \frac{N}{2}\sum_{j=d+1}^{p}\delta_j/\lambda_j$$

$$-\frac{N}{2}\log(\delta_1\cdots\delta_d) - \frac{N}{2}d \tag{4.5}$$

where the supremum is attained at $\delta_j = \lambda_j$ for $j = 1,2,\ldots,d$.

First, we assume that $\tau < k$. In this case

$$\sup_{\theta\in\Theta_k} L(\theta) = \sup_{\Phi(\tau,k)} L^*(\lambda_{\tau+1},\ldots,\lambda_p) \tag{4.6}$$

where $\sup_{\Phi(\tau,k)}$ indicates that the supremum is taken over $\lambda_{\tau+1} \geq \cdots \geq \lambda_k > 1$ and $\lambda_{k+1} = \cdots = \lambda_p = 1$. But

$$\sup_{\Phi(\tau,k)} L^*(\lambda_{\tau+1},\ldots,\lambda_p) = -\frac{N}{2}\sum_{i=1}^{\tau} \log \delta_i - \frac{N}{2}\tau$$

$$-\frac{N}{2}\sum_{i=k+1}^{p} \delta_i + \frac{N}{2}\sup_{\lambda_{\tau+1}\geq\cdots\geq\lambda_k>1}\{-\sum_{i=\tau+1}^{k} \log \lambda_i - \sum_{j=\tau+1}^{k} \delta_i/\lambda_i\} \tag{4.7}$$

Also, $\delta_i < 1$ and $\lambda_i > 1$ for $i = \tau+1,\ldots,k$. So,

$$-(\log \lambda_i + (\delta_i/\lambda_i)) < \delta_i \tag{4.8}$$

and the equality holds only when $\lambda_i = 1$. Since the above λ_i's can be arbitrarily approximated to one, we have

$$\sup_{\theta\in\Theta_k} L(\theta) = -\frac{N}{2}\sum_{i=1}^{\tau} \log\delta_i - N\tau/2 - \frac{N}{2}\sum_{i=\tau+1}^{p} \delta_i \tag{4.9}$$

when $\tau < k$. Next, let $\tau \geq k$. Then

$$\sup_{\theta\in\Theta_k} L(\theta) = -\frac{N}{2}\log(\delta_1\ldots\delta_k) - \frac{N}{2}k$$

$$+ \sup_{\lambda_{k+1}=\cdots=\lambda_p=1}\{-\frac{N}{2}\log(\lambda_{k+1}\ldots\lambda_p) - \frac{N}{2}\sum_{j=k+1}^{p}\delta_j/\lambda_j\}$$

$$= -\frac{N}{2}\sum_{i=1}^{k}\log\delta_i - \frac{N}{2}k - \frac{N}{2}\sum_{i=k+1}^{p}\delta_i\,. \tag{4.10}$$

Combining (4.9) and (4.10), we obtain

$$\sup_{\theta\in\Theta_k} L(\theta) = -\frac{N}{2}\sum_{i=1}^{p} \log\delta_i - \frac{Np}{2} + \frac{N}{2}\sum_{i=1+\min(\tau,k)}^{p}(\log\delta_i+1-\delta_i)\,. \tag{4.11}$$

But the supremum of $L(\theta)$ over the whole parametric space is given by

$$-\frac{N}{2}\sum_{i=1}^{p}\log\delta_i - \frac{Np}{2}$$

So, the logarithm of the likelihood ratio test statistic for testing Θ_k is given by

$$L_k = \frac{N}{2} \sum_{i=1+\min(\tau,k)}^{p} (\log\delta_i + 1 - \delta_i). \tag{4.12}$$

Now, let

$$\tilde{L}_k = \frac{N}{2} \sum_{i=k+1}^{p} (\log\delta_i + 1 - \delta_i). \tag{4.13}$$

We know from(3.7) that

$$\delta_i - \lambda_i = O\left(\left(\frac{\log\log N}{N}\right)^{1/2}\right) \quad \text{a.s.} \tag{4.14}$$

Suppose the true model is M_q. Then

$$\lambda_1 \geq \cdots \geq \lambda_q > \lambda_{q+1} = \cdots = \lambda_p = 1 . \tag{4.15}$$

From (4.14), we know with probability one, that $\delta_i > 1$ for $i = 1,2,\ldots,q$ and $\min(q,\tau) = q$ for large N. So, the statistics L_q and $\tilde{L}_q$ have the same distribution asymptotically. Here, we note that Anderson (1963) suggested to use $\tilde{L}_q$ as a statistic to test Θ_q and pointed out that the asymptotic distribution of $\tilde{L}_q$ is chi-square with $(p-q)(p-q+1)/2$ degrees of freedom. Rao (1983) pointed out that $\tilde{L}_q$ is not the LRT statistic.

We will now consider the problem of selecting one of the models $M_0, M_1, \ldots, M_{p-1}$ by using an information theoretic criterion. Let

$$G(k) = L_k - C_N k(2p-k+1)/2 \tag{4.16}$$

where C_N satisfies the following conditions

$$\text{(i)} \lim_{N\to\infty}(C_N/N) = 0, \qquad \text{(ii)} \lim_{N\to\infty}(C_N/\log\log N) = \infty .$$

We select the model $M_{\hat{q}}$ where $\hat{q}$ is chosen such that

$$G(\hat{q}) = \max_{0 \le k \le p-1} G(k). \tag{4.17}$$

We will now show that $\hat{q}$ is a consistent estimate of q.

THEOREM 4.1. If $N\hat{\Sigma}$ is distributed as central Wishart matrix with N degrees of freedom and $E(\hat{\Sigma}) = \Sigma$, then $\hat{q}$ is a strongly consistent estimate of q.

PROOF. Suppose that Θ_q is the true model and $k < q$. We have

$$G(q) - G(k) = L_q - L_k - C_N(q-k)(2p-k-q+1)/2. \tag{4.18}$$

As mentioned above, with probability one, we have for large N,

$$\delta_i > 1,\ i = 1,\dots,q, \text{ and } \min(q,\tau) = q. \tag{4.19}$$

Thus with probability one for large N,

$$\begin{aligned} L_q - L_k &= \frac{1}{2} N \sum_{i=q+1}^{p} (\log\delta_1 + 1 - \delta_i) - \frac{1}{2} N \sum_{i=k+1}^{p} (\log\delta_i + 1 - \delta_i) \\ &= -\frac{1}{2} N \sum_{i=k+1}^{q} (\log\delta_i + 1 - \delta_i) = \frac{1}{2} N\, W_N(q,k), \end{aligned}$$

where

$$W_N(q,k) = -\sum_{i=k+1}^{q} (\log\delta_i + 1 - \delta_i).$$

We have

$$\lim_{N\to\infty} W_N(q,k) \overset{a.s.}{=\!=} W(q,k) \equiv -\sum_{i=k+1}^{q} (\log\lambda_i + 1 - \lambda_i) > 0.$$

Hence, with probability one, we have for large N,

$$L_q - L_k > \frac{1}{4} N\, W(q,k),$$

and

$$G(q) - G(k) > 0. \tag{4.20}$$

Here we used the condition $\lim_{N\to\infty} C_N/N = 0$.

Now we assume that $k > q$. By (4.19) we have

$$|L_q - L_k| \le N \sum_{i=q+1}^{p} |\log\delta_i + 1 - \delta_i|.$$

Since $|\delta_i - 1| = O\left(\left(\frac{\log\log N}{N}\right)^{1/2}\right)$ a.s. for $i > q$, we can use Taylor's expansion, to get

$$|L_q - L_k| \le N \sum_{i=q+1}^{p} \frac{1}{2}(\delta_i - 1)^2(1+o(1)) \text{ a.s.}$$

$$= O(\log\log N) \quad \text{a.s.}$$

From $C_N/\log\log N \to \infty$, we see that with probability one, for large N,

$$G(q) - G(k) = O(\log\log N) + C_N(k-q)(2p-k-q+1)/2 > 0. \tag{4.21}$$

From (4.20) and (4.21), it follows that with probability one for large N,

$$\hat{q} = q.$$

Thus Theorem 4.1 is proved.

When the underlying distribution is complex multivariate normal, the proof for the consistency of the method goes along the same lines as in the real case.

5. DETERMINATION OF THE NUMBER OF SIGNALS WHEN THE UNDERLYING DISTRIBUTION IS ELLIPTICALLY SYMMETRIC

In this section, we discuss procedures for determination of the number of signals transmitted when the underlying distribution is real or complex elliptically symmetric. Here, we note that a random vector $\underset{\sim}{y}$ is said to to elliptically symmetric if its density is of the form

$$f(\underset{\sim}{y}) = |\Sigma|^{-1/2} g(\tfrac{1}{2}(\underset{\sim}{y}-\underset{\sim}{\mu})'\Sigma^{-1}(\underset{\sim}{y}-\underset{\sim}{\mu})) \tag{5.1}$$

where g is a non-increasing function in $[o,\infty)$. Multivariate normal and multivariate t distributions are special cases of the elliptically symmetric distributions. Kelker (1970) proposed the elliptically symmetric distributions and studied some of its properties. Krishnaiah and Lin (1984) proposed complex elliptically symmetric distribution and studied some of its properties. A complex random vector $\underset{\sim}{x} = \underset{\sim}{x}_1 + i\underset{\sim}{x}_2$ is said to be distributed as complex elliptically symmetric distribution if its density is of the form

$$f(\underset{\sim}{x}) = |\Sigma|^{-1} h((\underset{\sim}{x}-\underset{\sim}{\mu})'\Sigma^{-1}(\overline{\underset{\sim}{x}-\underset{\sim}{\mu}})) \tag{5.2}$$

where Σ is Hermitian, $\bar{\underset{\sim}{a}}$ denotes the complex conjugate of $\underset{\sim}{a}$, and h $(\cdot)$ is a non-increasing function in $[o,\infty)$. The covariance matrix of $(\underset{\sim}{x}_1', \underset{\sim}{x}_2')$ has the structure

$$\begin{pmatrix} \Sigma_1 & \Sigma_2 \\ -\Sigma_2 & \Sigma_1 \end{pmatrix}.$$

Complex multivariate normal considered by Wooding (1958) and Goodman (1963) and complex multivariate t distribution are special cases of the complex elliptically symmetric distribution. The density of the complex multivariate normal is known to be

$$f(\underset{\sim}{x}) = \frac{1}{\pi|\Sigma|} \exp\{-(\underset{\sim}{x}-\underset{\sim}{\mu})'\Sigma^{-1}(\overline{\underset{\sim}{x}-\underset{\sim}{\mu}})\}. \tag{5.3}$$

Now, consider the signal process $\underset{\sim}{x}(t)$ in (2.1) but assume that the joint density of $\underset{\sim}{x}_1 = \underset{\sim}{x}(t_1),\ldots,\underset{\sim}{x}_N = \underset{\sim}{x}(t_N)$ is

$$f(\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_N) = |\Sigma|^{-N} h(N \text{ tr } \Sigma^{-1}\hat{\Sigma}) \tag{5.4}$$

where $N\hat{\Sigma} = \sum_{j=1}^{N} \underset{\sim}{x}_j \bar{\underset{\sim}{x}}_j'$. Let $\lambda_1 \geq \ldots \geq \lambda_p$ be the eigenvalues of Σ and let $\ell_1 \geq \ldots \geq \ell_p$ denote the eigenvalues of $\hat{\Sigma}$. Also, let Θ_k denote the model in which

$$\lambda_1 \geq \ldots \geq \lambda_k > \lambda_{k+1} = \ldots = \lambda_p = \sigma^2 \tag{5.5}$$

where σ^2 is unknown. Let $f(\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_N|\Theta_k)$ denote the likelihood function under k-th model Θ_k. Also, let

$$L(\Theta_k) = \log f(\underset{\sim}{x}_1,\ldots,\underset{\sim}{x}_N|\Theta_k) \tag{5.6}$$

for $k = 0,1,\ldots,p-1$. We know that for given $\lambda_1,\ldots,\lambda_p$ the minimum of $\text{tr}\Sigma\hat{\Sigma}^{-1}$ is $\sum_{j=1}^{p} \lambda_i^{-1}\ell_i$ (see von Neumann (1937)). So,

$$\max_{\Theta_k} L(\Theta_k) = \max_{\Theta_k}\{-N\sum_{j=1}^{p} \log \lambda_j + \log h(N\sum_{j=1}^{p} \lambda_j^{-1}\ell_j)\} \tag{5.7}$$

where the maximum is taken subject to (5.5). Suppose $h(t)$ has a continuous derivative $h'(t)$ on $[0,\infty)$ and the equation

$$Nph(y) = y|h'(y)| \tag{5.8}$$

has a unique solution $y = Np/\gamma_h$. Then, the above maximum is reached at

$$\frac{\lambda_1}{\ell_1} = \dots = \frac{\lambda_k}{\ell_k} = \frac{\sigma^2}{\{\frac{1}{p-k}(\ell_{k+1}+\dots+\ell_p)\}} = \gamma_h \tag{5.9}$$

and

$$\max_{\Theta_k} L(\Theta_k) = -(Np/2)\log\gamma_h + \log h(Np/\gamma_h) - N\sum_{i=1}^{p}\log\ell_i + G_1(k) \tag{5.10}$$

where

$$G_1(k) = N\log\left[\prod_{i=k+1}^{p}\ell_i/\left(\frac{1}{p-k}\sum_{i=k+1}^{p}\ell_i\right)^{p-k}\right] . \tag{5.11}$$

Under the conditions of Theorem 3.1. we observe that, for $k < q$,

$$\lim_{N\to\infty}\frac{1}{N}(G_1(k)-G_1(k-1)) = W(k,k-1) > 0 \quad \text{a.s.} \tag{5.12}$$

and for $k > q(k \le p-1)$

$$\frac{1}{N}(G_1(k)-G_1(k-1)) = O\left(\frac{\log\log N}{N}\right) \quad \text{a.s.} \tag{5.13}$$

where

$$W(k,k-1) = (p-k+1)[\log\left(\frac{1}{(p-k+1)}(\lambda_{k+1}+\dots+\lambda_p\right)$$

$$-\frac{1}{(p-k+1)}\log\lambda_k - \frac{(p-k)}{(p-k+1)}\log\left(\frac{1}{(p-k)}(\lambda_{k+1}+\dots+\lambda_p)\right) . \tag{5.14}$$

We see that $G_1(k)$ is non-decreasing function of k for $k\in\{0,1,\dots,p-1\}$. If

we draw the points $(0, G_1(0)), (1,G_1(1)), \ldots,(p-1,G_1(p-1))$ in the Descartes coordinate plane, and construct a polygonal line with these points as its p vertexes, then $G_1(k)-G_1(k-1)$ is just the slope of the k-th segment. Suppose that q is the true number of signals. For convenience we temporarily assume $q > 0$. As shown in (5.12) and (5.13), we can assert with probability one that, for large N,

$$G_1(k) - G_1(k-1) \geq C_1 N \text{ for } k \leq q \tag{5.15}$$

and

$$G_1(k) - G_1(k-1) = O(\log\log N) \quad \text{for } q < k \leq p-1, \tag{5.16}$$

where $C_1 > 0$ is a constant. Thus we see that, the slope $G_1(k) - G_1(k-1)$ has a significant change for $k \leq q$ and $q < k \leq p-1$, and the true value q is just the largest k for which $G_1(k) - G_1(k-1) > C_N$, where C_N satisfies the following conditions:

$$\lim_{N\to\infty}(C_N/N) = 0 \qquad \lim_{N\to\infty}(C_N/\log\log N) = \infty . \tag{5.17}$$

If we put $G_1(-1) = -\infty$, then the same is true for $q = 0$. Motivated by (5.15) and (5.16), we estimate the number of signals q with $\hat{q}$ where

$$\hat{q} = \max\{k \leq p-1: G_1(k) - G_1(k-1) > C_N\}. \tag{5.18}$$

Under the conditions of Theorem 3.1, we can show that $\hat{q}$ is a consistent estimate of q by following the same lines as in Section 3.

In general, we do not know whether the conditions of Theorem 3.1 are

satisfied. In these cases, we make the following assumptions:

$$\text{(i)} \lim_{N\to\infty} \frac{1}{N}(G_1(k)-G_1(k-1)) \overset{a.s}{=} 0 \quad \text{for } k > q$$

$$\text{(ii)} \lim_{N\to\infty} \frac{1}{N}(G_1(k)-G_1(k-1)) \overset{a.s}{=} W(k,k-1) > 0 \text{ for } k \le q \tag{5.18}$$

where we denote $\lambda_0 = \infty$ for convenience. In this case, we need to assume that the smallest non-zero eigenvalue of $A\psi\bar{A}'$ is distinguishable from σ^2, namely, the ratio of signal intensity to that of noise can be detected by the sensor. We assume that $(\lambda_q-\sigma^2)/\sigma^2 \ge \varepsilon > 0$ and ε is known for the given receiver. In this case, we estimate q with $\hat{q}$ where $\hat{q}$ is chosen such that

$$\hat{q} = \max\{k \le p-1: G_1(k) - G_1(k-1) > \frac{\mu}{2} N\}, \tag{5.19}$$

where we denote $G_1(-1) = -\infty$ for convenience. Also,

$$\mu = \min_{0\le k\le p-1} (p-k+1)\{\log(\frac{1}{p-k+1} + \frac{p-k}{p-k+1}\delta) - \frac{p-k}{p-k+1}\log\delta\} > 0, \tag{5.20}$$

and

$$\delta = 1 - \frac{\varepsilon}{p(1+\varepsilon)} \cdot \tag{5.21}$$

We now establish the strong consistency of $\hat{q}$. To prove this, we write

$$\alpha_k = \frac{1}{p-k+1}, \quad \beta_k = \frac{p-k}{p-k+1},$$

$$A_k = \frac{1}{p-k} \sum_{i=k+1}^{p} \lambda_i/\lambda_k, \ \lambda_0 = \infty .$$

Suppose that q is the true number of signals and $k \le q$. Then

$$\lim_{N\to\infty} \frac{1}{N}(G_1(k)-G_1(k-1)) = W(k,k-1) \text{ a.s.} \tag{5.22}$$

$$W(k,k-1) = (p-k+1)\{\log(\alpha_k+\beta_k A_k) - \beta_k \log A_k\} > 0.$$

Consider $f_k(x) = \log(\alpha_k+\beta_k x) - \beta_k \log x$ for x (0,1]. We have

$$f_k'(x) = -\alpha_k\beta_k(1-x)/x(\alpha_k+\beta_k x) < 0,\ 0 < x \leq 1,$$

so that $f_k(x)$ is a decreasing function on (0,1]. But if $\lambda_q \geq (1+\varepsilon)\sigma^2$, then for $0 \leq k \leq q-1$,

$$A_k \leq \frac{p-q}{p-k}\,\frac{1}{1+\varepsilon} + \frac{q-k}{p-k} \leq 1 - \frac{1}{p-k}\,\frac{\varepsilon}{1+\varepsilon} \leq 1 - \frac{\varepsilon}{p(1+\varepsilon)} = \delta$$

and

$$A_q = \sigma^2/\lambda_q \leq \frac{1}{1+\varepsilon} < \delta\ .$$

Thus for $0 \leq k \leq q$,

$$W(k,k-1) \geq \mu\ . \tag{5.23}$$

From (5.22) and (5.23), it follows that, with probability one for large N,

$$G_1(k) - G_1(k-1) > \frac{\mu}{2} N,\quad k \leq q. \tag{5.24}$$

On the other hand, if $q < k \leq p-1$,

$$\lim_{N\to\infty} \frac{1}{N}(G_1(k)) - G_1(k-1)) = 0 \quad \text{a.s.} \tag{5.25}$$

So with probability one, for large N,

$$G_1(k) - G_1(k-1) < \frac{\mu}{2} N \text{ for } q < k \leq p-1. \tag{5.26}$$

Thus from (5.24) and (5.26) it follows, with probability one, for large N,

$$\hat{q} = q. \tag{5.27}$$

and the assertion is proved.

REFERENCES

Akaike, H. (1972). Information theory and an extension of the maximum likelihood principle. In Proceedings of the Second International Symposium on Information Theory, Supp. to Problems of Control and Information Theory, pp. 267-281.

Anderson, T.W. (1963). Asymptotic theory for principal component analysis. Annals of Mathematical Statistics, 34, 122-148.

Bai, Z. D. (1984). A note on asymptotic joint distribution of the eigenvalues of a noncentral multivariate F matrix. Tech. Report No. 84-49. Center for Multivariate Analysis, University of Pittsburgh.

Bartlett, M.S. (1954). A note on the multiplying factors for various χ^2 approximations. Journal of the Royal Statistical Society, Ser. B, 16, 296-298.

Goodman, N.R. (1963). Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction). Annals of Mathematical Statistics, 34, 152-176.

Hannan, E.J. and Quinn, B.G. (1979). The determination of the order of an autoregression. Journal of the Royal Statistical Society, Ser. B, 41, 190-195.

Kelker, D. (1970). Distribution theory of spherical distribution and a location scale parameter generalization. Sankhya, Ser. A, 43, 419-430.

Krishnaiah, P.R. (1976). Some recent developments on complex multivariate distributions. Journal of Multivariate Analysis, 6, 1-30.

Krishnaiah, P.R. and Lin, J. (1984). Complex elliptical distributions. Technical Report No. 84-19. Center for Multivariate Analysis, University of Pittsburgh.

Krishnaiah, P.R. and Schuurmann, F.J. Computations of Complex Multivariate Distributions (in preparation). To be published by North-Holland publishing Company.

Krishnaiah, P.R. and Waikar, V.B. (1971). Simultaneous tests for equality of latent roots against certain alternatives, I. Annals of the Institute of Statistical Mathematics, 23, 451-468.

Krishnaiah, P.R. and Waikar, V.B. (1972). Simultaneous tests for equality of latent roots against certain alternatives, II. Annals of the Institute of Statistical Mathematics, 24, 81-85.

Liggett, W.S. (1973). Passive sonar: fitting models to multiple time series, In Signal Processing (J.W. Griffith et al, editors) pp. 327-345. Academic Press, New York.

Rao, C.R. (1983). Likelihood ratio tests for relationships between two covariance matrices. In Studies in Econometrics, Time Series and Multivariate Statistics (T.Amemiya, S.Karlin and L. Goodman, editors). Academic Press, Inc. New York.

Reznik, M.K. (1968). The law of the iterated logarithm for some classes of stationary processes. Theory of Probability and Its Applications, 8, 606-621.

Rissanen, J. (1978). Modeling by shortest data description. Automatica, 14, 465-471.

Schmidt, R.O. (1979). Multiple emitter location and signal parameter estimation. In Proc. RADC Spectrum Estimation Workshop, Rome.

Schwartz, G. (1978). Estimating the dimension of a model. Annals of Statistics, 6, 461-464.

Stout, W.F. (1974). Almost Sure Convergence. Academic Press. New York.

Von Neumann, J. (1937). Some matrix-inequalities and metrization of metric space. Tomsk Univ. Rev., 1, 286-300.

Wax, M. and Kailath, T. (1984). Determination of the number of signals by information theoretic criteria. IEEE Transactions on Acoustics, Speech, and Signal Processing, ASSP-33

Wax, M., Shan, T.J., and Kailath, T. (1984). Spatio-temporal spectral analysis by eigenstructure methods. IEEE Transactions on Acoustics, Speech and Signal Processing, ASSP-32, 817-827.

Whalen, A.D. (1971). Detection of Signals in Noise. Academic Press. New York.

Wooding, R.A. (1956). The multivariate distribution of complex normal variables. Biometrika, 43, 212-215.

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE	READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 86-0039 ~~85-37~~ 2. GOVT ACCESSION NO. AD-A164211	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) On detection of number of signals in presence of white noise	5. TYPE OF REPORT & PERIOD COVERED Technical - October 1985
	6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) L. C. Zhao, P. R. Krishnaiah, and Z. D. Bai	8. CONTRACT OR GRANT NUMBER(s) ~~N00014-85-K-0292 (Navy)~~ F49620-85-C-0008 (~~Air Force~~)
9. PERFORMING ORGANIZATION NAME AND ADDRESS Center for Multivariate Analysis 515 Thackeray Hall University of Pittsburgh, Pittsburgh, PA 15260	10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 2304 / A5
11. CONTROLLING OFFICE NAME AND ADDRESS ~~Office of Naval Research and~~ Air Force Office of Scientific Research / NM Bolling AFB DC 20332	12. REPORT DATE October 1985
	13. NUMBER OF PAGES 26
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office)	15. SECURITY CLASS. (of this report) Unclassified
	15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Elliptically symmetric distribution, information theory, model selection, principal component analysis, signal detection and white noise.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

In this paper, the authors propose procedures for the detection of the number of signals in presence of Gaussian white noise. The methods used fall within the framework of the model selection procedures using an information theoretic criterion. The strong consistency of the estimates of the number of signals, under different situations, is established. Extensions of the results are also discussed to the case when the noise is not necessarily Gaussian.

DD FORM 1 JAN 73 1473

Unclassified

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

END

FILMED

3-86

DTIC